Solutions - Assignment 6 - 2022111024 - Q1.ipynb

Q1. Prove that X^TX is invertible if and only if the columns of X are linearly independent.

Soln. Hence, we are to prove two things:

($\Rightarrow$) X^TX is invertible if the columns of X are linearly independent.

($\Leftarrow$) The columns of X are linearly independent if ~~th~~ X^TX is invertible, these prove the if and only if statement given.

($\Rightarrow$) Forward implication proof:

X has linearly independent columns. To show, X^TX is invertible.

We know, when ~~x~~ all columns of X are linearly independent, $X\vec{y}=\vec{b}$ ~~$A\vec{x}=\vec{b}$~~ has a solution $\forall \vec{b}$ ($\because$ n independent vectors of ~~$n\times n$ dimensi~~ X with n columns span the n-dimensional vector space, so a solution $\forall b$ will exist).

Thus, ~~$A\vec{x}=\vec{b}$~~ $X\vec{y}=\vec{b}$ has a solution ~~$\vec{x}$~~ $\vec{y}$ ~~if~~ since X has linearly independent columns.

For X to be invertible, there must exist ~~A^{-1}~~ X^{-1} s.t. ~~$AA^{-1}=A^{-1}A=I$~~ $XX^{-1}=X^{-1}X=I$.

Let $\{\vec{e_1}, \vec{e_2}, \dots \vec{e_n}\}$ be the columns of the identity matrix.

$$XX^{-1}\cancel{AA^{-1}} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{e_1} & \vec{e_2} & \dots & \vec{e_n} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

We can show matrix multiplication as a series of matrix-vector multiplications such that,

$$X\cancel{A}\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{a_1}^* & \vec{a_2}^* & \dots & \vec{a_n} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ X\cancel{A}\vec{a_1}^* & X\cancel{A}\vec{a_2}^* & \dots & X\cancel{A}\vec{a_n} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{e_1} & \vec{e_2} & \dots & \vec{e_n} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$X\vec{y} = \vec{b}$

∵ ~~$A\vec{x} = \vec{b}$~~ always has a soln, we can calculate each $\vec{a}_i^*$ from the equation $X\vec{a}_i^* = \vec{e}_i \ \forall i \in \{1, 2, \dots n\}$ and we can guarantee a soln. to it will exist.

Thus, solving $\forall \vec{a}_i^*$ $(i \in \{1, 2, .. n\})$ we ~~get a~~ can create a matrix whose column vectors are the solutions to $\vec{a}_i^*$ for each column i, such that

$$X \cancel{A} \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ a_1^* & a_2^* & \dots & a_n^* \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = I.$$

∵ ~~A~~ soln to $\vec{a}_i^*$ for each $\vec{a}_i^*$ always exists, we can always find such as a matrix, giving us X^{-1} ~~A^{-1}~~. Thus X ~~A~~ is invertible.

Thm ~~we know~~

$$(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$$
$$A^T (A^{-1})^T = (A^{-1}A)^T = I^T = I.$$

$A^T B^T = (BA)^T$ and $I^T = I$.

So, A^T always has an inverse $(A^T)^{-1}$ given by $(A^{-1})^T$ whenever A has an inverse A^{-1}.

Thus, A^T is invertible if A is invertible.

∵ X is invertible, X^T is invertible. Now, product of 2 invertible matrices is always invertible as can be shown by,

Thm $C = AB$, $\exists A^{-1}$ s.t. $AA^{-1} = A^{-1}A = I$, $BB^{-1} = B^{-1}B = I$.

Multiplying B^{-1} on both sides,
$$CB^{-1} = ABB^{-1}$$
$$\Rightarrow CB^{-1} = AI = A$$
Multiplying A^{-1} on both sides,
$$CB^{-1}A^{-1} = AA^{-1} = I.$$

Multiplying A^{-1} on both sides
$$A^{-1}C = A^{-1}AB \cancel{A}$$
$$\Rightarrow A^{-1}C = IB = B$$
Multiplying ~~A~~ B^{-1} on both sides,
$$B^{-1}A^{-1}C = B^{-1}B = I$$

∴ $\exists B^{-1}A^{-1}$ ~~s.t.~~ $= C^{-1}$ s.t. $CC^{-1} = C^{-1}C = I$.

Hence, X^TX is invertible.

X has linearly independent columns $\Rightarrow X^TX$ is invertible.

($\Leftarrow$) Backward implication proof:

X^TX is invertible. To show, columns of X are linearly independent.

To prove by contradiction, assume the columns of X are ~~lined~~ not linearly independent. $\therefore \exists c_1, c_2, \ldots c_n$ s.t not all of them are zero s.t. $c_1x_1 + c_2x_2 + \ldots + c_nx_n = 0$ where x_i represents the i^{th} column of X, i.e. $X = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ x_1 & x_2 & \ldots & x_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$

Thus, we can rewrite the eqn. as,

$$X \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = 0 \quad \text{where } \exists \text{ some } k \in \{[1,n]\} \text{ s.t. } c_k \neq 0$$

Multiplying X^T on both sides,

$$\therefore X^T X \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = X^T . 0$$

$$\Rightarrow X^TX \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = 0$$

~~$\therefore$~~ $\because X^TX$ is invertible and $\exists$ ~~c_k~~ k s.t. $c_k \neq 0$, we can say that $\exists$. a non zero vector $\vec{c}$ such that $(X^TX)\vec{c} = 0$. Thus, $\exists$ non-zero vector in the null space of X^TX.

This contracts ~~that~~ the fact that X^TX is invertible. Thus, the columns of X are linearly independent, $\because$ only then ~~it~~ will ~~there~~ ~~a~~ $\exists\ \vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$, s.t. $(X^TX)^{-1}\vec{c} = 0$

$\because$ for linearly independent x_i, ~~x_2~~ ~~$c_i \neq 0$ for an~~ $c_i = 0\ \forall i$, thus $\vec{c} = \vec{0}$, which is the only vector belonging to the null space of X^TX, an invertible matrix.

$\therefore X^TX$ is invertible $\Rightarrow$ columns of X are linearly independent.

$\therefore X^TX$ is invertible $\Longleftrightarrow$ colums of X are linearly independent.

Proved

Q2 Find the solution for β if the matrix X is decomposed as $X = QR$ using QR decomposition where Q is orthogonal & R is an upper triangular matrix.

Soln. Using QR decomposition, we find $X = QR$ where Q is orthogonal & R is upper tr.

We know that β is given by $X\beta = Y$.

Multiplying X^T on both sides,

$$\Rightarrow X^T(X\beta) = X^T Y$$
$$\Rightarrow (X^T X)\beta = X^T Y$$

Substituting $X = QR$,

$$[(QR)^T(QR)]\beta = (QR)^T Y$$

$$\Rightarrow [(R^T Q^T)(QR)]\beta = (R^T Q^T) Y$$

$$\Rightarrow [R^T (Q^T Q) R]\beta = R^T Q^T Y$$

$\because$ Q is orthogonal, $Q^T Q = I$.

$$\Rightarrow R^T I R \beta = R^T Q^T Y$$
$$\Rightarrow R^T R \beta = R^T Q^T Y$$

$\because R$ is upper triangular, R^T is invertible (diagonal elements nonzero)

$\Rightarrow \therefore R^T$ is invertible [proved in Q1]

$\therefore (R^T)^{-1}$ exists s.t. $(R^T)^{-1} \cdot R^T = I$

Multiplying $(R^T)^{-1}$ on both sides,

$$(R^T)^{-1} \cdot R^T R \beta = (R^T)^{-1} \cdot R^T \cdot Q^T Y$$
$$\Rightarrow I R\beta = I Q^T Y$$
$$\Rightarrow R\beta = Q^T Y.$$

$\boxed{R\beta = Q^T Y}$ Let $Q^T Y = W$, $\therefore R\beta = W$, we do not need to take the inverse of R to solve for β with $\beta = R^{-1} W$ since R is an upper triangular matrix.

~~Instead, if W_i is the W can be written as its~~

Instead, we solve for β_i $[i \in \{1, \dots n\}]$ using W_i $[i \in \{1, \dots n\}]$ using upper Δr elements of R given by R_{ij}, $[i, j \in \{1, \dots n\}]$.

$$\beta_n = z_n / R_{nn}$$
$$\beta_{n-1} = (z_{n-1} - R_{(n-1)n}\beta_n) / R_{n-1\,n-1}$$
$$\vdots$$
$$\beta_1 = z_1 - R_{12}\beta_2 - \dots R_{1n}\beta_n) / R_{11}$$

Thus we have calculated $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_n \end{bmatrix}$

Q3. Find the solution for β if the matrix X is decomposed into $X = UDV^T$ using singular value decomposition, where U, V are orthogonal matrices & D is diagonal matrix.

$X = UDV^T$ $\qquad$ $X\beta = Y$

~~or, from normal equation,~~

$$\cancel{(X^TX^{-1})^{-1}X^T}$$

Or, from Moore-Penrose equation given to us,

$$\therefore (X^TX)^{-1} X^TY = \beta$$

$$\Rightarrow \beta = ((UDV^T)^T(UDV^T))^{-1}(UDV^T)^TY$$
$$= ((VD^TU^T)(UDV^T))^{-1}(UDV^T)^TY$$
$$= (VD^T(U^TU)DV^T)^{-1}(UDV^T)^TY$$
$$= (VD^TIDV^T)^{-1}(UDV^T)^TY$$

$$=(VD^TDV^T)^{-1}(UDV)^TY$$

∵ Diagonal matrix D is symmetric ∵ it has no non-diagonal elements to be inverted. Thus, $D^T = D$.

$$= (VD^2V^T)^{-1}(VDU^T)Y$$

$$= \left[(V^T)^{-1}D^{-2}V^{-1}\right](VDU^T)Y$$

$$= \left((V^T)^T D^{-2} V^{-1}\right)(VDU^T)Y$$

Since, V is orthogonal, V^T is orthogonal. $V^{-1} = V^T$. Similarly $(V^T)^{-1} = (V^T)^T$

We need not bother about calculating D^{-1} ∵ it is a diagonal matrix and its inverse is given by the diagonal matrix formed ~~the~~ by the multiplicative inverse of each diagonal element in D, as is ~~giv~~ shown by,

if $D = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_n \end{bmatrix}$, $D^{-1} = \begin{bmatrix} 1/\sigma_1 & 0 & 0 & 0 \\ 0 & 1/\sigma_2 & \vdots & \vdots \\ 0 & 0 & \cdots & 1/\sigma_n \end{bmatrix}$

$$\Rightarrow \beta = (VD^{-2}V^{-1})(VDU^T)Y$$

$$\Rightarrow \beta = (VD^{-2}(V^{-1}V)DU^T)Y$$

$$\Rightarrow \beta = (VD^{-2}IDU^T)Y$$

$$\Rightarrow \beta = (VD^{-2}DU^T)Y$$

$\Rightarrow \boxed{\beta = (VD^{-1}\cancel{V}U^T)Y}$, we simply need to ~~f~~ transpose V^T and U on decomposing X and find the multiplicative inverses of the diagonal elements of D to form D^{-1} making ~~it~~ ~~co~~ computation of β much faster.